#2   OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/970,287

DATE: 10/24/2001
TIME: 16:28:57

Input Set : **A:\10147611.app**
Output Set: **N:\CRF3\10242001\I970287.raw**

```
 3 <110> APPLICANT: GLUCKSMANN, Maria A.
 4        MEYERS, Rachel
 5        KAPELLER-LIBERMANN, Rosana
 6        SILOS-SANTIAGO, Inmaculada
 8 <120> TITLE OF INVENTION: 22437, A NOVEL HUMAN SULFATASE AND USES THEREFOR
10 <130> FILE REFERENCE: 10147-61U1
12 <140> CURRENT APPLICATION NUMBER: US/09/970,287 OK
13 <141> CURRENT FILING DATE: 2001-10-03
15 <150> PRIOR APPLICATION NUMBER: US 60/257,082
16 <151> PRIOR FILING DATE: 2000-12-21
18 <160> NUMBER OF SEQ ID NOS: 12
20 <170> SOFTWARE: PatentIn Ver. 2.1
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 3513
24 <212> TYPE: DNA
25 <213> ORGANISM: Homo sapiens
27 <400> SEQUENCE: 1
28 cgcgtccggg cagattcacg tcgtttccag ccaagtggac ctgatcgatg gccctcctga 60
29 atttatcacg atatttgatt tattagcgat gcccctggt ttgtgtgtta cgcacacaca 120
30 cgtgcacaca aggctctggc tcgcttccct ccctcgtttc cagctcctgg gcgaatccca 180
31 catctgtttc aactctccgc cgagggcgag caggagcgag agtgtgtcga atctgcgagt 240
32 gaagagggac gagggaaaag aaacaaagcc acagacgcaa cttgagactc ccgcatccca 300
33 aaagaagcac cagatcagca aaaaagaag atgggccccc cgagcctcgt gctgtgcttg 360
34 ctgtccgcaa ctgtgttctc cctgctgggt ggaagctcgg ccttcctgtc gcaccaccgc 420
35 ctgaaaggca ggtttcagag ggaccgcagg aacatccgcc ccaacatcat cctggtgctg 480
36 acggacgacc aggatgtgga gctgggttcc atgcaggtga tgaacaagac ccggcgcatc 540
37 atggagcagg gcgggacgca cttcatcaac gccttcgtga ccacacccat gtgctgcccc 600
38 tcacgctcct ccatcctcac cggcaagtac gtccacaacc acaacaccta caccaacaat 660
39 gagaactgct cctcgccctc ctggcaggca cagcacgaga gccgcacctt tgccgtgtac 720
40 ctcaatagca ctggctaccg gacagctttc ttcgggaagt atcttaatga atacaacggc 780
41 tcctacgtgc cacccggctg gaaggagtgg gtcggactcc ttaaaaactc ccgcttttat 840
42 aactacacgc tgtgtcggaa cggggtgaaa gagaagcacg gctccgacta ctccaaggat 900
43 tacctcacag acctcatcac caatgacagc gtgagcttct ccgcacgtc caagaagatg 960
44 tacccgcaca ggccagtcct catggtcatc agccatgcag cccccacgg ccctgaggat 1020
45 tcagccccac aatattcacg cctcttccca aacgcatctc agcacatcac gccgagctac 1080
46 aactacgcgc ccaacccgga caaacactgg atcatgcgct acacggggcc catgaagccc 1140
47 atccacatgg aattcaccaa catgctccag cggaagcgct tgcagaccct catgtcggtg 1200
48 gacgactcca tggagacgat ttacaacatg ctggttgaga cgggcgagct ggacaacacg 1260
49 tacatcgtat acaccgccga ccacggttac cacatcggcc agtttggcct ggtgaaaggg 1320
50 aaatccatgc catatgagtt tgacatcagg gtcccgttct acgtgagggg ccccaacgtg 1380
51 gaagccggct gtctgaatcc ccacatcgtc ctcaacattg acctggcccc caccatcctg 1440
52 gacattgcag gcctggacat acctgcggat atggacggga atccatcct caagctgctg 1500
53 gacacggagc ggccggtgaa tcggtttcac ttgaaaaaga agatgagggt ctggcgggac 1560
54 tccttcttgg tggagagagg caagctgcta cacaagagag acaatgacaa ggtggacgcc 1620
55 caggaggaga actttctgcc caagtaccag cgtgtgaagg acctgtgtca gcgtgctgag 1680
56 taccagacgg cgtgtgagca gctgggacag aagtggcagt gtgtggagga cgccacgggg 1740
57 aagctgaagc tgcataagtg caagggcccc atgcggctgg cggcagcag agccctctcc 1800
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/970,287

DATE: 10/24/2001
TIME: 16:28:57

Input Set : A:\10147611.app
Output Set: N:\CRF3\10242001\I970287.raw

```
58 aacctcgtgc ccaagtacta cgggcagggc agcgaggcct gcacctgtga cagcgggggac 1860
59 tacaagctca gcctggccgg acgccggaaa aaactcttca agaagaagta caaggccagc 1920
60 tatgtccgca gtcgctccat ccgctcagtg gccatcgagg tggacggcag ggtgtaccac 1980
61 gtaggcctgg gtgatgccgc ccagccccga aacctcacca agcggcactg gccaggggcc 2040
62 cctgaggacc aagatgacaa ggatggtggg gacttcagtg gcactggagg ccttcccgac 2100
63 tactcagccg ccaaccccat taaagtgaca catcggtgct acatcctaga gaacgacaca 2160
64 gtccagtgtg acctggacct gtacaagtcc ctgcaggcct ggaaagacca caagctgcac 2220
65 atcgaccacg agattgaaac cctgcagaac aaaattaaga acctgaggga agtccgaggt 2280
66 cacctgaaga aaaagcggcc agaagaatgt gactgtcaca aaatcagcta ccacacccag 2340
67 cacaaaggcc gcctcaagca cagaggctcc agtctgcatc ctttcaggaa gggcctgcaa 2400
68 gagaaggaca aggtgtggct gttgcgggag cagaagcgca agaagaaact ccgcaagctg 2460
69 ctcaagcgcc tgcagaacaa cgacacgtgc agcatgccag gcctcacgtg cttcacccac 2520
70 gacaaccagc actggcagac ggcgcctttc tggacactgg ggcctttctg tgcctgcacc 2580
71 agcgccaaca ataacacgta ctggtgcatg aggaccatca atgagactca caatttcctc 2640
72 ttctgtgaat ttgcaactgg cttcctagag tactttgatc tcaacacaga cccctaccag 2700
73 ctgatgaatg cagtgaacac actggacagg gatgtcctca accagctaca cgtacagctc 2760
74 atggagctga ggagctgcaa gggttacaag cagtgtaacc cccggactcg aaacatggac 2820
75 ctgggactta agatggagg aagctatgag caatacaggc agtttcagcg tcgaaagtgg 2880
76 ccagaaatga agagaccttc ttccaaatca ctgggacaac tgtgggaagg ctgggaaggt 2940
77 taagaaacaa cagaggtgga cctccaaaaa catagaggca tcacctgact gcacaggcaa 3000
78 tgaaaaacca tgtgggtgat ttccagcaga cctgtgctat tggccaggag gcctgagaaa 3060
79 gcaagcacgc actctcagtc aacatgacag attctggagg ataaccagca ggagcagaga 3120
80 taacttcagg aagtccattt ttgccctgc ttttgctttg gattatacct caccagctgc 3180
81 acaaaatgca tttttcgta tcaaaaagtc accactaacc ctcccccaga agctcacaaa 3240
82 ggaaaacgga gagagcgagc gagagagatt tccttggaaa tttctcccaa gggcgaaagt 3300
83 cattggaatt tttaaatcat aggggaaaag cagtcctgtt ctaaatcctc ttattctttt 3360
84 ggtttgtcac aaagaaggaa ctaagaagca ggacagaggc aacgtggaga ggctgaaaac 3420
85 agtgcagaga cgtttgacaa tgagtcagta gcacaaaaga gatgacattt acctagcact 3480
                                                                     3513
86 ataaaccctg gttgcctctg aagaaactgc ctt
```

```
89 <210> SEQ ID NO: 2
90 <211> LENGTH: 870
91 <212> TYPE: PRT
92 <213> ORGANISM: Homo sapiens
94 <400> SEQUENCE: 2
95 Met Gly Pro Pro Ser Leu Val Leu Cys Leu Leu Ser Ala Thr Val Phe
96   1               5                  10                  15
98 Ser Leu Leu Gly Gly Ser Ser Ala Phe Leu Ser His His Arg Leu Lys
99              20                  25                  30
101 Gly Arg Phe Gln Arg Asp Arg Arg Asn Ile Arg Pro Asn Ile Ile Leu
102              35                  40                  45
104 Val Leu Thr Asp Asp Gln Asp Val Glu Leu Gly Ser Met Gln Val Met
105          50                  55                  60
107 Asn Lys Thr Arg Arg Ile Met Glu Gln Gly Gly Thr His Phe Ile Asn
108  65                  70                  75                  80
110 Ala Phe Val Thr Thr Pro Met Cys Cys Pro Ser Arg Ser Ser Ile Leu
111              85                  90                  95
113 Thr Gly Lys Tyr Val His Asn His Asn Thr Tyr Thr Asn Asn Glu Asn
114             100                 105                 110
116 Cys Ser Ser Pro Ser Trp Gln Ala Gln His Glu Ser Arg Thr Phe Ala
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/970,287

Input Set : A:\10147611.app
Output Set: N:\CRF3\10242001\I970287.raw

```
117              115                      120                      125
119 Val Tyr Leu Asn Ser Thr Gly Tyr Arg Thr Ala Phe Phe Gly Lys Tyr
120     130                      135                      140
122 Leu Asn Glu Tyr Asn Gly Ser Tyr Val Pro Pro Gly Trp Lys Glu Trp
123 145                      150                      155                      160
125 Val Gly Leu Leu Lys Asn Ser Arg Phe Tyr Asn Tyr Thr Leu Cys Arg
126                      165                      170                      175
128 Asn Gly Val Lys Glu Lys His Gly Ser Asp Tyr Ser Lys Asp Tyr Leu
129                      180                      185                      190
131 Thr Asp Leu Ile Thr Asn Asp Ser Val Ser Phe Phe Arg Thr Ser Lys
132              195                      200                      205
134 Lys Met Tyr Pro His Arg Pro Val Leu Met Val Ile Ser His Ala Ala
135     210                      215                      220
137 Pro His Gly Pro Glu Asp Ser Ala Pro Gln Tyr Ser Arg Leu Phe Pro
138 225                      230                      235                      240
140 Asn Ala Ser Gln His Ile Thr Pro Ser Tyr Asn Tyr Ala Pro Asn Pro
141                      245                      250                      255
143 Asp Lys His Trp Ile Met Arg Tyr Thr Gly Pro Met Lys Pro Ile His
144              260                      265                      270
146 Met Glu Phe Thr Asn Met Leu Gln Arg Lys Arg Leu Gln Thr Leu Met
147     275                      280                      285
149 Ser Val Asp Asp Ser Met Glu Thr Ile Tyr Asn Met Leu Val Glu Thr
150     290                      295                      300
152 Gly Glu Leu Asp Asn Thr Tyr Ile Val Tyr Thr Ala Asp His Gly Tyr
153 305                      310                      315                      320
155 His Ile Gly Gln Phe Gly Leu Val Lys Gly Lys Ser Met Pro Tyr Glu
156                      325                      330                      335
158 Phe Asp Ile Arg Val Pro Phe Tyr Val Arg Gly Pro Asn Val Glu Ala
159              340                      345                      350
161 Gly Cys Leu Asn Pro His Ile Val Leu Asn Ile Asp Leu Ala Pro Thr
162     355                      360                      365
164 Ile Leu Asp Ile Ala Gly Leu Asp Ile Pro Ala Asp Met Asp Gly Lys
165     370                      375                      380
167 Ser Ile Leu Lys Leu Leu Asp Thr Glu Arg Pro Val Asn Arg Phe His
168 385                      390                      395                      400
170 Leu Lys Lys Lys Met Arg Val Trp Arg Asp Ser Phe Leu Val Glu Arg
171              405                      410                      415
173 Gly Lys Leu Leu His Lys Arg Asp Asn Asp Lys Val Asp Ala Gln Glu
174              420                      425                      430
176 Glu Asn Phe Leu Pro Lys Tyr Gln Arg Val Lys Asp Leu Cys Gln Arg
177              435                      440                      445
179 Ala Glu Tyr Gln Thr Ala Cys Glu Gln Leu Gly Gln Lys Trp Gln Cys
180     450                      455                      460
182 Val Glu Asp Ala Thr Gly Lys Leu Lys Leu His Lys Cys Lys Gly Pro
183 465                      470                      475                      480
185 Met Arg Leu Gly Gly Ser Arg Ala Leu Ser Asn Leu Val Pro Lys Tyr
186              485                      490                      495
188 Tyr Gly Gln Gly Ser Glu Ala Cys Thr Cys Asp Ser Gly Asp Tyr Lys
189              500                      505                      510
```

```
191 Leu Ser Leu Ala Gly Arg Arg Lys Lys Leu Phe Lys Lys Lys Tyr Lys
192         515                     520                 525
194 Ala Ser Tyr Val Arg Ser Arg Ser Ile Arg Ser Val Ala Ile Glu Val
195         530                     535                 540
197 Asp Gly Arg Val Tyr His Val Gly Leu Gly Asp Ala Ala Gln Pro Arg
198 545                     550                     555                 560
200 Asn Leu Thr Lys Arg His Trp Pro Gly Ala Pro Glu Asp Gln Asp Asp
201                 565                     570                     575
203 Lys Asp Gly Gly Asp Phe Ser Gly Thr Gly Gly Leu Pro Asp Tyr Ser
204             580                     585                     590
206 Ala Ala Asn Pro Ile Lys Val Thr His Arg Cys Tyr Ile Leu Glu Asn
207             595                     600                     605
209 Asp Thr Val Gln Cys Asp Leu Asp Leu Tyr Lys Ser Leu Gln Ala Trp
210         610                     615                     620
212 Lys Asp His Lys Leu His Ile Asp His Glu Ile Glu Thr Leu Gln Asn
213 625                     630                     635                     640
215 Lys Ile Lys Asn Leu Arg Glu Val Arg Gly His Leu Lys Lys Lys Arg
216             645                     650                     655
218 Pro Glu Glu Cys Asp Cys His Lys Ile Ser Tyr His Thr Gln His Lys
219             660                     665                     670
221 Gly Arg Leu Lys His Arg Gly Ser Ser Leu His Pro Phe Arg Lys Gly
222         675                     680                     685
224 Leu Gln Glu Lys Asp Lys Val Trp Leu Leu Arg Glu Gln Lys Arg Lys
225         690                     695                     700
227 Lys Lys Leu Arg Lys Leu Leu Lys Arg Leu Gln Asn Asn Asp Thr Cys
228 705                     710                     715                     720
230 Ser Met Pro Gly Leu Thr Cys Phe Thr His Asp Asn Gln His Trp Gln
231                 725                     730                     735
233 Thr Ala Pro Phe Trp Thr Leu Gly Pro Phe Cys Ala Cys Thr Ser Ala
234             740                     745                     750
236 Asn Asn Asn Thr Tyr Trp Cys Met Arg Thr Ile Asn Glu Thr His Asn
237             755                     760                     765
239 Phe Leu Phe Cys Glu Phe Ala Thr Gly Phe Leu Glu Tyr Phe Asp Leu
240         770                     775                     780
242 Asn Thr Asp Pro Tyr Gln Leu Met Asn Ala Val Asn Thr Leu Asp Arg
243 785                     790                     795                     800
245 Asp Val Leu Asn Gln Leu His Val Gln Leu Met Glu Leu Arg Ser Cys
246                 805                     810                     815
248 Lys Gly Tyr Lys Gln Cys Asn Pro Arg Thr Arg Asn Met Asp Leu Gly
249             820                     825                     830
251 Leu Lys Asp Gly Gly Ser Tyr Glu Gln Tyr Arg Gln Phe Gln Arg Arg
252             835                     840                     845
254 Lys Trp Pro Glu Met Lys Arg Pro Ser Ser Lys Ser Leu Gly Gln Leu
255         850                     855                     860
257 Trp Glu Gly Trp Glu Gly
258 865                     870
261 <210> SEQ ID NO: 3
262 <211> LENGTH: 2610
263 <212> TYPE: DNA
```

```
264 <213> ORGANISM: Homo sapiens
266 <400> SEQUENCE: 3
267 atgggcccc cgagcctcgt gctgtgcttg ctgtccgcaa ctgtgttctc cctgctgggt 60
268 ggaagctcgg ccttcctgtc gcaccaccgc ctgaaaggca ggtttcagag ggaccgcagg 120
269 aacatccgcc ccaacatcat cctggtgctg acggacgacc aggatgtgga gctgggttcc 180
270 atgcaggtga tgaacaagac ccggcgcatc atggagcagg gcgggacgca cttcatcaac 240
271 gccttcgtga ccacacccat gtgctgcccc tcacgctcct ccatcctcac cggcaagtac 300
272 gtccacaacc acaacaccta caccaacaat gagaactgct cctcgccctc ctggcaggca 360
273 cagcacgaga gccgcacctt tgccgtgtac ctcaatagca ctggctaccg gacagctttc 420
274 ttcgggaagt atcttaatga atacaacggc tcctacgtgc cacccggctg gaaggagtgg 480
275 gtcggactcc ttaaaaactc ccgctttat aactacacgc tgtgtcggaa cggggtgaaa 540
276 gagaagcacg gctccgacta ctccaaggat tacctcacag acctcatcac caatgacagc 600
277 gtgagcttct tccgcacgtc caagaagatg tacccgcaca ggccagtcct catggtcatc 660
278 agccatgcag cccccacgg ccctgaggat tcagccccac aatattcacg cctcttccca 720
279 aacgcatctc agcacatcac gccgagctac aactacgcgc ccaacccgga caaacactgg 780
280 atcatgcgct acacggggcc catgaagccc atccacatgg aattcaccaa catgctccag 840
281 cggaagcgct tgcagaccct catgtcggtg gacgactcca tggagacgat ttacaacatg 900
282 ctggttgaga cgggcgagct ggacaacacg tacatcgtat acaccgccga ccacggttac 960
283 cacatcggcc agtttggcct ggtgaaaggg aaatccatgc catatgagtt tgacatcagg 1020
284 gtcccgttct acgtgagggg ccccaacgtg gaagccggct gtctgaatcc ccacatcgtc 1080
285 ctcaacattg acctggcccc caccatcctg gacattgcag gcctggacat acctgcggat 1140
286 atggacggga aatccatcct caagctgctg tccttcttgg tggagagagg caagctgcta 1200
287 ttgaaaaaga agatgagggt ctggcgggac tccttcttgg tggagagagg caagctgcta 1260
288 cacaagagag acaatgacaa ggtggacgcc caggaggaga actttctgcc caagtaccag 1320
289 cgtgtgaagg acctgtgtca gcgtgctgag taccagacgg cgtgtgagca gctgggacag 1380
290 aagtggcagt gtgtggagga cgccacgggg aagctgaagc tgcataagtg caagggcccc 1440
291 atgcggctgg gcggcagcag agccctctcc aacctcgtgc ccaagtacta cgggcagggc 1500
292 agcgaggcct gcacctgtga cagcggggac tacaagctca gcctggccgg acgccggaaa 1560
293 aaactcttca agaagaagta caaggccagc tatgtccgca gtcgctccat ccgctcagtg 1620
294 gccatcgagg tggacggcag ggtgtaccac gtaggcctgg gtgatgccgc ccagccccga 1680
295 aacctcacca agcggcactg gccaggggcc cctgaggacc aagatgacaa ggatggtggg 1740
296 gacttcagtg gcactggagg ccttcccgac tactcagccg ccaacccat taaagtgaca 1800
297 catcggtgct acatcctaga gaacgacaca gtccagtgtg acctggacct gtacaagtcc 1860
298 ctgcaggcct ggaaagacca caagctgcac atcgaccacg agattgaaac cctgcagaac 1920
299 aaaattaaga acctgaggga agtccgaggt cacctgaaga aaaagcggcc agaagaatgt 1980
300 gactgtcaca aaatcagcta ccacacccag cacaaaggcc gcctcaagca cagaggctcc 2040
301 agtctgcatc ctttcaggaa gggcctgcaa gagaaggaca aggtgtggct gttgcgggag 2100
302 cagaagcgca agaagaaact ccgcaagctg ctcaagcgcc tgcagaacaa cgacacgtgc 2160
303 agcatgccag gcctcacgtg cttcacccac gacaaccagc actggcagac ggcgcctttc 2220
304 tggacactgg ggcctttctg tgcctgcacc agcgccaaca taacacgta ctggtgcatg 2280
305 aggaccatca atgagactca caatttcctc ttctgtgaat ttgcaactgg cttcctagag 2340
306 tactttgatc tcaacacaga cccctaccag ctgatgaatg cagtgaacac actggacagg 2400
307 gatgtcctca accagctaca cgtacagctc atggagctga ggagctgcaa gggttacaag 2460
308 cagtgtaacc cccggactcg aaacatggac ctgggactta agatggagg aagctatgag 2520
309 caatacaggc agtttcagcg tcgaaagtgg ccagaaatga agagaccttc ttccaaatca 2580
310 ctgggacaac tgtgggaagg ctgggaaggt                                      2610
313 <210> SEQ ID NO: 4
314 <400> SEQUENCE: 4
315 000
```

W--> 314
W--> 315

OR

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/970,287

DATE: 10/24/2001
TIME: 16:28:58

Input Set : A:\10147611.app
Output Set: N:\CRF3\10242001\I970287.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application Number
L:314 M:283 W: Missing Blank Line separator, <400> field identifier
L:315 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (4) SEQUENCE:
L:319 M:283 W: Missing Blank Line separator, <400> field identifier
L:320 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (5) SEQUENCE:
L:324 M:283 W: Missing Blank Line separator, <400> field identifier
L:325 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (6) SEQUENCE:
L:329 M:283 W: Missing Blank Line separator, <400> field identifier
L:330 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (7) SEQUENCE:
L:334 M:283 W: Missing Blank Line separator, <400> field identifier
L:335 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (8) SEQUENCE:
L:339 M:283 W: Missing Blank Line separator, <400> field identifier
L:340 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (9) SEQUENCE:
L:344 M:283 W: Missing Blank Line separator, <400> field identifier
L:345 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (10) SEQUENCE:

10/24/01